

Acquisition de la variation phonologique et recueil de corpus d'interactions naturelles parents-enfants : nouvelle méthode, nouveaux enjeux

Chabanal Damien, lrl, Université Blaise Pascal, Clermont-Ferrand

Liégeois Loïc, lrl, Université Blaise Pascal, Clermont-Ferrand

Chanier Thierry, lrl, Université Blaise Pascal, Clermont-Ferrand

Mots-clés : liaison ; schwa ; corpus denses ; discours adressé à l'enfant ; effets de fréquence

Keywords : liaison ; schwa ; dense corpora ; child directed speech ; frequency effects

Cet article propose d'étudier le rapport que peuvent entretenir la recherche sur l'acquisition de la variation phonologique et les méthodes d'analyse à partir de corpus denses d'interactions parents-enfant. L'objectif de ce travail est double. Dans un premier temps il s'agit d'illustrer, à partir de résultats issus de travaux précédents, les avantages et les limites d'une analyse de l'acquisition de la variation phonologique menée sur des corpus denses dans le cadre des modèles basés sur l'usage. Nous exposerons notamment comment la mesure des effets de fréquence dans l'input sur la mise en place progressive de schémas de variation phonologique doit être envisagée. Dans un second temps, l'objectif de notre travail est de montrer comment l'analyse des interactions parents-enfant recueillies en situation naturelle remet en cause une analyse purement phonologique de l'acquisition de la variation phonologique, et plus particulièrement de l'élision du schwa dans les monosyllabes.

This article investigates analysis methods of phonological variation. More precisely, it addresses the relation between research on the acquisition of phonological variation and methods of corpus-based analyses of interactions between parents and their children. The objective of this research is twofold. First, we attempt to illustrate, based on concrete examples from our previous studies' findings, the advantages and the limitations of a corpus-based analysis of the acquisition of phonological variation within the framework of usage-based theories. In fact we will attempt to demonstrate the necessity of measuring frequency effects in the input on the progressive construction of phonological variation schemas. Secondly, we would like to show how the analysis of parental -child interactions collected in natural everyday situations challenges the mere phonological analysis of the acquisition of phonological variation, and especially of schwa elision in monosyllabic words.

## Introduction

La maîtrise de la variabilité des usages linguistiques est selon nous empirique et émergente. Une des hypothèses majeures concernant la construction d'une grammaire hétérogène ou de patrons phonologiques généraux chez l'enfant repose sur le contact avec l'usage (Tomasello, 2003 ; Chevrot, Dugua et Fayol, 2009). La fréquence de rencontres des formes variables, directement issues de l'input, permettrait de donner des indices stables aux jeunes sujets sur la manière dont sont construits des schémas généraux de variation. Ainsi, plus certains indices seraient présents dans l'environnement, plus ils faciliteraient la construction de schémas, sur la base de la mémorisation de constructions lexicales fréquentes ayant des caractéristiques grammaticales ou phonologiques communes. Cette hypothèse d'une acquisition de la variation phonologique basée sur l'usage réclame, pour être vérifiée, la prise en compte de données quantitativement importantes tout autant chez l'enfant que dans son bain linguistique. Les corpus denses, à l'œuvre dans cette réflexion, sont au cœur de notre travail sur la question de l'acquisition de la variation phonologique. Nous présentons ici, à travers les corpus denses du projet ALIPE (Liégeois, Chanier et Chabanal, 2014) la méthode de recueil envisagée pour l'étude de l'acquisition de la liaison et de l'élision. Nous faisons part également des nouveaux enjeux que cette approche soulève, notamment sur les questions de l'influence de l'adresse du discours parental et plus généralement du poids des contraintes sociolinguistiques par rapport à des contraintes proprement phonologiques.

## Les corpus d'interactions parents-enfant

L'ère moderne des travaux sur l'acquisition du langage, que Ingram (1989) fait débiter avec la publication de Taine (1876), a vu se multiplier et se systématiser les tâches de recueil de données spontanées. Au départ, celles-ci étaient recueillies sous forme de notes compilées dans un journal (ou *parental diary*) et ont alors servi de base aux premières études s'appuyant sur des échantillons relativement importants d'énoncés enfantins produits en situation naturelle d'interaction. Les études ne se focalisent pas principalement sur le développement linguistique, mais l'intègre à des analyses portant sur le développement cognitif général de l'enfant (Darwin, 1877 ; Taine, 1877). Puis progressivement certains recueils vont se concentrer principalement sur le langage des enfants (Grégoire, 1948 ; Leopold, 1939). Par souci pratique, c'est généralement le parent du sujet étudié qui se charge de la prise de notes. Cet aspect méthodologique présente un certain avantage, dans le sens où le développement de l'enfant est généralement décrit précisément au niveau de l'axe temporel. En effet, un parent apparaît le mieux placé pour témoigner de l'apparition d'un nouveau mot dans le lexique de l'enfant ou d'un nouveau type de construction syntaxique.

Cependant, le recueil de notes se révèle critiquable à deux niveaux. Tout d'abord, la méthode mise en place induit nécessairement un problème de subjectivité puisque la source des données primaires n'est accessible qu'au moment de la prise de notes, sans retour en arrière possible. De ce fait, les observations récoltées (commentaires, transcription orthographiques ou phonétiques) sont basées sur la seule perception instantanée de l'enquêteur.

Une deuxième critique peut être émise au sujet de la spécificité de ces corpus. En effet, étant donné qu'il est impossible de capturer tous les éléments de la situation de recueil, le relevé effectué par l'enquêteur est conditionné par une méthodologie et des hypothèses de recherche définies en amont. Il apparaît ainsi impossible de croiser les données des différents journaux dans le but de mener des études quantitatives et générales sur le développement linguistique des enfants (Morgenstern et Parisse, 2007). D'autres facteurs comme le temps passé avec l'enfant et la particularité des situations d'interaction (Dugua, 2006), font que leurs études ne sont pas généralisables et sont souvent considérées comme une description du système phonologique d'un enfant particulier capturé à un moment précis (Demuth, 2011 ; Fikkert, 2000). C'est d'ailleurs l'une des caractéristiques majeures des *parental diaries*, et donc des analyses qui en découlent : elles sont centrées sur les performances de l'enfant. Ainsi, jusque dans les années 1970, le discours adressé à l'enfant (désormais DAE) ne fait l'objet d'aucune attention particulière (Cross, 1979).

## Des cadres méthodologiques et théoriques nouveaux

L'absence d'étude centrée sur les caractéristiques du DAE avant les années 1970 n'est pas uniquement due à des problèmes méthodologiques. En effet, les principaux cadres théoriques qui ont influencé les travaux des chercheurs ne considéraient pas le DAE comme un élément moteur du développement linguistique des enfants. Alors qu'aucune étude sérieuse, qualitative ou quantitative, n'a été réalisée sur des productions de parents en interaction avec leur enfant, le DAE est décrit dans la littérature scientifique comme un input « pauvre », partiellement agrammatical et similaire au discours entre adultes (Chomsky, 1965). L'input reçu par les enfants, ponctué de faux départs, de pauses et « d'erreurs » (conjugaison et accord par exemple), ne serait pas assez riche pour leur offrir l'ensemble des indices leur permettant de construire leur grammaire. Ces arguments ont servi (et servent encore) le courant générativiste qui considère la « pauvreté du stimulus » comme « un argument majeur qui soutient le postulat d'une grammaire finie innée » (Wauquier-Gravelines, 2005, p. 332). Si l'on considère que l'environnement de l'enfant ne joue pas un rôle clef dans le processus d'acquisition du langage, il n'y a donc aucune raison de l'étudier. Ainsi les travaux du cadre générativiste ont davantage cherché à décrire ce qu'était la compétence linguistique, la grammaire d'un locuteur, plutôt que d'observer les processus de son développement.

Les années 1960-1970 vont marquer un tournant majeur au niveau théorique et méthodologique. Le développement des cadres socio-interactionnistes et cognitifs mettent en avant le rôle de l'interaction de l'enfant avec son environnement dans les processus de développement cognitif. Le principe d'innéité des capacités langagières de l'enfant est sérieusement remis en question, tout comme l'autonomie du langage par rapport aux autres compétences cognitives (Bruner, 1977, 1983). Le cadre de l'interaction est étudié de près, et des études émergent sur le déroulement des échanges parent-enfant (Bruner, 1974), le paralinguistique (Halliday, 1973) et la nature du DAE (Snow, 1972). De l'avis de Garnica (1977), cette prise en compte du cadre de l'interaction et du DAE va avoir deux conséquences majeures :

*One possible consequence is that greater emphasis is placed on the mother-child interaction and the contribution of this experience to the child's learning of language [...]. Another consequence is that child's role becomes more important since the child can be seen as a more active participant in the language-learning process. (p. 64)*

Les deux points soulevés par Garnica trouvent aujourd'hui un écho dans les modèles basés sur l'usage dont le postulat de base est que la compétence linguistique se développe à partir des événements d'usage, à savoir les énoncés entendus et produits en contexte (Tomasello, 2003). Les chercheurs travaillant dans ce cadre émettent la nécessité de récolter un nouveau type de données : les corpus denses. En effet, les données les plus répandues jusqu'alors sont des corpus longitudinaux, extrêmement utiles dans le but d'étudier le développement linguistique d'un enfant sur une période relativement longue (de une à plusieurs années). Cependant, cette méthodologie a des limites. En effet, en se basant sur une à deux heures d'enregistrement par mois, les corpus longitudinaux permettent de capturer seulement entre 1% et 1,5% du discours infantin, en considérant que l'enfant est éveillé et en situation de communication environ dix heures par jour (Diessel, 2008 ; Rowland, Fletcher et Freudenthal, 2008 ; Tomasello et Stahl, 2004).

Certains chercheurs se lancent alors dans la récolte d'enregistrements beaucoup plus fréquents sur une courte période dans le but de capturer, en fonction de la méthodologie mise en place, jusqu'à 20% de la totalité des énoncés produits par l'enfant sur cette période (Lieven et Behrens, 2012). Par exemple, en enregistrant l'enfant en interaction environ une heure par jour pendant cinq jours, les chercheurs disposent alors d'environ 10% de ses productions (Diessel, 2008 ; Rowland, Fletcher et Freudenthal, 2008). Même si cette méthodologie de recueil des données permet de saisir une faible proportion des énoncés produits et entendus par l'enfant, les corpus denses, comparés aux corpus longitudinaux classiques composés d'enregistrements relativement courts et espacés dans le temps, sont beaucoup plus efficaces lorsqu'il s'agit de relever des phénomènes peu fréquents ou ayant lieu pendant une courte période de l'acquisition (Rowland, Fletcher et Freudenthal, 2008).

L'utilisation de corpus denses règle donc certains problèmes rencontrés avec les corpus longitudinaux classiques. Ils fournissent des données différentes de ces derniers et permettent non seulement de « relever les premiers usages d'une structure linguistique particulière, mais également d'évaluer si une structure se développe spécifiquement et lexicalement ou d'une façon plus générale » (Lieven et Behrens, 2012, p. 228. Notre traduction).

## Variation phonétique et phonologique en discours adressé à l'enfant

Les caractéristiques phonétiques et phonologiques du discours adressé à l'enfant ont la plupart du temps été comparées à celles relevées en discours adressé à l'adulte (désormais DAA) dans le but de confirmer les propriétés prosodiques relevées « à la main » telles que le ralentissement du débit ou la hauteur tonale plus élevée (Biersack, Kempe et Knapton, 2005 ; Fergusson, 1964 ; Fernald, Taeschner, Dunn, Papousek, De Boysson-Bardies et Fukui, 1989 ; Schaeffler et Kempe, 2007). Les propriétés segmentales du DAE, contrairement aux unités suprasegmentales, ont reçu peu d'attention de la part des chercheurs, surtout en ce qui concerne la variation de leur production (Dilley, Millet, Mcauley et Bergeson, 2013 ; Foulkes, Docherty et Watt, 2005 ; Watt, Docherty et Foulkes, 2003). Cependant, quelques spécificités peuvent être dégagées.

Dans une étude portant sur la variation de production du /t/ en anglais, Foulkes et ses collègues (Foulkes, Docherty et Watt, 2005) ont recueilli des interactions entre quarante mères et leur enfant âgé de 2;0 ans à 4;0 ans. En ce qui concerne les données en DAA, les auteurs se sont appuyés sur des conversations informelles entre locutrices de la même région et de la même classe sociale. Il s'avère qu'en position médiane, alors que les variantes standard du /t/ représentent seulement 10% des occurrences en DAA, ce taux passe à 59% en DAE. La variante glottale de la consonne, réalisée dans 90% des cas en DAA, ne représente quant à elle que 36% des contextes en DAE.

Ainsi, il apparaît que même si les enfants sont davantage exposés à la variante standard que les adultes, ces derniers sont également confrontés à davantage de variation. Ces résultats peuvent être mis en relation avec ceux obtenus par Andreassen (2013) concernant les schwas présents en syllabe initiale des polysyllabes. Dans son travail de doctorat, Andreassen a notamment observé le comportement du schwa en DAA et en DAE dans une variété suisse du français. Ses données révèlent également un suremploi de la variante standard (avec production du schwa) en DAE par rapport à ses données du DAA tirées du corpus « Phonologie du Français Contemporain » (Durand, Laks et Lyche, 2002). Il s'avère par exemple que, lorsque le schwa est entouré de deux fricatives (comme dans *faisant*), la voyelle est élidée dans 99% des cas en DAA, contre 45% en DAE. Dans ce contexte précis, il apparaît donc que les enfants sont exposés à une variation beaucoup plus importante que les adultes.

## Acquisition de la variation phonologique et corpus denses

L'utilisation de corpus denses recueillis en milieu naturel, au domicile familial, nous paraît intéressante à plusieurs niveaux. Premièrement, ce type de données permet une comparaison directe du DAE et du DAA. En effet, deux contextes d'interaction sont généralement recueillis dans ce genre de situation : des interactions parent(s)-enfant et père-mère. Ainsi, il n'est pas nécessaire de procéder à une nouvelle récolte de données pour comparer les deux types d'interaction. Cette méthode, en plus d'être avantageuse au niveau méthodologique, permet également d'éviter les biais inhérents au fait que ce sont les productions de locuteurs différents qui sont comparées entre elles. Ce paramètre devient déterminant lorsqu'il s'agit d'étudier des phénomènes variables pour lesquels les facteurs sociolinguistiques jouent un rôle déterminant (comme pour la liaison et l'élision par exemple).

Ensuite, l'étude de corpus denses recueillis en situation naturelle d'interaction favorise la prise en compte de l'ensemble des niveaux entrant en jeu dans les phénomènes de variation : lexique, syntaxe, situation d'énonciation etc. L'analyse de contextes nombreux et variés permet ainsi de mesurer les différents facteurs conditionnant la variation. La réalisation de la liaison variable peut ainsi être étudiée en fonction de la nature du mot liaisonnant, du contexte morphosyntaxique de production ou du type d'interaction (conversation au cours d'un repas ou lors d'un jeu de rôle entre parents et enfants par exemple). Enfin, ce type de données est également utile dans le but de tester les scénarios d'acquisition proposés dans un cadre *usage-based*. En effet, puisque les événements d'usage sont considérés comme déterminants au cours du processus d'acquisition, il paraît primordial de vérifier si les données issues des corpus corroborent les résultats souvent obtenus en situation d'expérimentation. Le modèle exemplariste basé sur l'usage que nous défendons est conditionné par des effets de fréquence d'usage (Chevrot, Chabanal et Dugua 2007 ; Chevrot, Dugua et Fayol, 2009 ; Chevrot, Dugua, Harnois-Delpiano, Siccardi et Spinelli, 2013). Ce modèle repose sur le fait que l'enfant mémoriserait en premier lieu des expressions figées fréquentes en input constituées des deux mots concernés par la liaison et de la consonne de liaison (le Mot1 *des*, la consonne /z/ et le Mot2 *ours* par exemple). Lors d'étapes ultérieures, des comportements de type analogiques, toujours conditionnés par la fréquence d'usage en input et donc sur des éléments concrets, engendreraient des mécanismes abstraits mais sans que les généralisations effacent les traces concrètes emmagasinées dans notre lexique. Ainsi, le schéma modéliserait l'encodage lexical de la variation phonologique sur des bases empiriques liées à la fréquence lexicale tout en fonctionnant sur des schémas plus abstraits composés de slots et/ou d'éléments lexicalisés, du type *Un + X*. Nos précédents travaux se sont donc focalisés sur la nature du phonème qui suit le monosyllabe : cette annotation permet de rendre compte du phonème produit après le monosyllabe, dans sa forme pleine (*le chat*) ou élidé (*l(e) chat*). Sauf les rares exceptions où le schwa est produit alors que le monosyllabe précède un mot à initiale vocalique<sup>1</sup>, les phonèmes qui suivent sont essentiellement des consonnes.

Le tableau ci-dessous résume l'ensemble des annotations qui portent sur les contextes d'élision et de maintien du schwa des monosyllabes en prenant comme exemple l'énoncé *il pourrait me fâcher*. En plus des annotations de base concernant les mots de gauche et de droite, la (non) réalisation du schwa et l'adresse du discours, nous retrouvons l'ensemble des annotations effectuées a posteriori et décrites plus haut.

Contexte gauche	Type mot de gauche	Polysyllabe
	Structure syllabique de gauche	CVCV
	Mot1	<i>pourrait</i>

<sup>1</sup> Ces contextes sont peu présents dans nos données et sont réalisés par les enfants. Il s'agit d'énoncés du type « le ours ».

	Phonème qui précède	/e/
<b>Élision</b>	Monosyllabe étudié	<i>me</i>
	Élision	Oui
	Adresse du discours	DAE
<b>Contexte droit</b>	Type mot de droite	Polysyllabe
	Structure syllabique de droite	CVCV
	Mot2	<i>fâcher</i>
	Phonème qui suit	/f/

Tableau 1 : annotations des corpus sur le schwa dans un monosyllabe

## L'acquisition de la variation phonologique : qu'est ce que les corpus d'interactions naturelles parents-enfant (ne) peuvent (pas) nous apprendre ?

L'objectif de cette partie porte sur les liens pouvant exister entre le lexique parental et le lexique enfantin et plus particulièrement entre les liaisons produites par les parents et leur enfant.

Chabanal et Liégeois (2014), Liégeois, Chabanal et Chanier (2011) et Liégeois (2014) ont ainsi pu démontrer que les données du DAE issues de leurs corpus denses allaient dans le sens du scénario de l'acquisition de la liaison proposé par Chevrot et ses collègues (2007, 2009, 2013). À partir d'un corpus dense regroupant les interactions entre une fillette âgée de 40 mois (Prune) et ses parents, Chabanal et Liégeois (2011) ont par exemple cherché à vérifier le poids de la fréquence lexicale des exemplaires de Mots2 en input sur la production des erreurs du type « un z-oiseau » ou « deux n-oreilles ». Les exemplaires de Mots2 correspondent aux différentes variantes que mémoriserait l'enfant lorsqu'il découpe la chaîne sonore en suivant le schéma CV (Chevrot, Chabanal et Dugua, 2007). Ainsi, en ce qui concerne le mot « ours », l'enfant pourrait mémoriser dans son lexique mental les exemplaires « n-ours » et « z-ours » à partir d'énoncés comme « un ours » et « des ours » par exemple. En clair, les auteurs se sont posé la question suivante : est-ce que l'enfant entendant et produisant plus fréquemment un exemplaire (par exemple « n-oiseau ») l'élira plus souvent en contexte d'erreur (par exemple « le beau n-oiseau ») ? Les résultats, obtenus à partir de l'analyse de 110 erreurs de liaisons relevées dans les productions de Prune, montrent dans la majorité des cas (2/3) une tendance chez la fillette à employer l'exemplaire de Mots2 le plus souvent produit par elle-même et par ses parents en contexte de liaison juste. L'effet direct de la fréquence lexicale dans l'acquisition de la liaison dans ses premières étapes semble donc réel.

## Méthodologies de recueil, de structuration et de traitement des données

Le corpus ALIPE (Liégeois, Chanier et Chabanal, 2014) regroupe trois corpus denses recueillis en situation naturelle d'interaction. Trois enfants (deux fillettes et un garçon) ont été enregistrés sur deux ou trois temps à leur domicile. Nous avons confié à chacun des couples de parents un enregistreur numérique en leur donnant pour seule consigne d'enregistrer leur enfant lors de situations propices aux interactions comme le bain, le repas ou au cours de moments de jeu. Cette méthodologie comporte plusieurs avantages. Premièrement, elle minimise les biais en n'imposant pas la présence d'un observateur extérieur à la famille lors des sessions d'enregistrement. Ensuite, elle permet de recueillir des données produites en situation naturelle en faisant varier les contextes de production. Ceci a ainsi favorisé la récolte d'énoncés adultes produits en direction de l'enfant ainsi que des énoncés issus de conversations entre les parents. Pour cette étude, nous nous focaliserons sur les productions des deux fillettes et de leurs parents. La couverture de chacun des deux corpus est exposée dans les tableaux ci-dessous.

	<b>Corpus Salomé</b>	
	Premier temps de recueil	Deuxième temps de recueil
<b>Âge de la fillette</b>	2;4 ans	3;0 ans
<b>Durée des enregistrements</b>	5 heures	4 heure 55 minutes
<b>Nombre de graphies transcrites</b>	29 788	26 102
<b>Nombre de monosyllabes</b>	3 241	2 798

Tableau 2 : présentation du corpus de Salomé

	<b>Corpus Prune</b>	
	Premier temps de recueil	Deuxième temps de recueil
<b>Âge de la fillette</b>	3;4 ans	5;4 ans
<b>Durée des enregistrements</b>	5 heures	4 heures 3 minutes
<b>Nombre de graphies transcrites</b>	33 568	25 491
<b>Nombre de monosyllabes</b>	2 511	2 270

Tableau 3 : présentation du corpus de Prune

Une fois recueillis, les enregistrements ont été transcrits, annotés et structurés dans deux formats distincts : le format XML-TEI (Text Encoding Initiative) et le format CHAT (Codes for the Human Analysis of Transcripts). En effet, ces deux formats de structuration apparaissent complémentaires (Liégeois, 2013). Alors que le premier offre une solution de structuration standard, extensible, pérenne et interopérable, le second permet l'analyse du corpus à l'aide du logiciel associé CLAN (Computerized Language ANalysis ; MacWhinney, 2000). De plus, le corpus ALIPE viendra prochainement enrichir la banque de corpus du projet CHILDES (MacWhinney et Snow, 1985 ; MacWhinney, 2000).

### Annotation de la variation phonologique

Alors que la transcription du corpus ALIPE a été effectuée au niveau orthographique, nous avons procédé à plusieurs annotations afin de mettre en lumière la variation phonologique. Ainsi, chaque contexte de liaison potentielle a été annoté dans le but de rendre compte de l'ensemble des paramètres utiles à notre étude. Nous avons ainsi renseigné, pour chaque contexte de liaison :



- Le contexte syntaxique de liaison, au moyen de la méthodologie mise en place par Chevrot, Chabanal et Dugua (2007). Cette annotation permet de renseigner si le contexte de liaison se trouve entre un déterminant ou un nom (contexte A), un pronom personnel et un verbe (contexte B), etc.
- La consonne attendue si la liaison est réalisée.
- La consonne de liaison effectivement réalisée par le locuteur, ou l'absence de réalisation.
- Le caractère variable, catégorique ou erratique du contexte de liaison.

En ce qui concerne le schwa, son élision est annotée au moyen des conventions définies par le format CHAT. Selon ces conventions, un segment élide est renseigné entre parenthèses. Cette annotation nous permet donc de rendre compte de l'élision du schwa (*j'ai pris l(e) train*) ou de sa réalisation (*j'ai pris le train*).

Pour réaliser cette étude, nous avons également procédé à un ensemble d'annotations effectué a posteriori à partir des contextes de maintien ou d'élision du schwa extraits de nos données.

## L'annotation a posteriori des critères phonologiques de réalisation ou d'élision du schwa

Afin de mesurer l'effet du contexte phonologique sur la (non) réalisation du schwa, nous avons procédé à une série d'annotations afin de rendre compte, le plus précisément possible, de la structure phonologique des contextes droit et gauche. Ces annotations correspondent à plusieurs critères :

- Le nombre de syllabes du mot suivant et précédant le monosyllabe étudié : nous distinguons, avec cette annotation, les polysyllabes des monosyllabes. Cette annotation n'est pas réalisée à partir de la forme orthographique de base mais en fonction de la forme réalisée. Ainsi, si le mot *chemin* est produit *ch(e)min*, il sera catégorisé comme monosyllabe. De plus, il est possible qu'aucun mot ne soit présent avant le monosyllabe si celui-ci a été produit en début d'énoncé.
- La structure de la syllabe qui suit et précède le monosyllabe étudié : ce critère d'annotation permet de rendre compte de la structure des syllabes qui entourent le monosyllabe étudié. Il permet, entre autre, de distinguer les syllabe ouvertes (Consonne-Voyelle), fermées (Voyelle-Consonne) ou à attaque complexe (CCV).

La partie qui suit expose les avantages et les limites actuelles de l'utilisation d'un corpus dense d'interactions parents-enfant pour l'étude de la variation phonologique et de son acquisition. La question repose principalement sur le fait de savoir quels apports peuvent représenter les corpus denses pour la sociolinguistique développementale.

## Fréquence des unités lexicales : intérêts et limites

La fréquence est le facteur fondamental des modèles basés sur l'usage servant de cadre à nos analyses. Comme l'écrit Loiseau (2011), ces modèles « ont mis au premier plan la fréquence comme facteur expliquant l'émergence et la systématisation des faits linguistiques à partir de la répétition de l'usage » (p. 59). Les processus de grammaticalisation et de syntactisation seraient directement liés à des dispositifs d'automatisation sous jacents à la fréquence. Dans cette configuration, plus la fréquence d'une unité est élevée, plus elle conduira à un meilleur ancrage. L'automatisation est directement en lien avec la fréquence. Le fait de pouvoir relever de grandes quantités de contextes variables identiques chez les parents et les enfants assure théoriquement de pouvoir observer l'influence de l'input parental sur le développement phonologique.

## Fréquence et acquisition de la liaison

Pour vérifier la force de la fréquence de l'input parental sur la production de liaisons catégoriques ou variables, nous avons observé et calculé un certain nombre de corrélations visant à attester de l'effet de l'input parental sur la variation. Pour ce faire, nous avons récolté trois types de données :

- des collocations lexicales Mot1-Mot2 équivalentes chez les parents et leur enfant (par exemple : c'est (Mot1) à (Mot2), petit ami...),
- des mêmes Mots1 avec des Mots 2 différents (par exemple, l'ensemble des liaisons du type : c'est (Mot1) + X à initiale vocalique (Mot2), petit + X à initiale vocalique...),
- des mêmes Mots2 avec des Mots 1 différents (par exemple X (Mot1) + a (Mot2), X + ont...).

Le ratio entre le nombre total de contextes de liaisons chez nos sujets (parents et enfants) et le nombre de contextes lexicaux identiques s'est avéré important. À titre d'exemple, dans le corpus regroupant les interactions entre Prune et ses parents, nous présentons les données pour le contexte « B » répertoriant les liaisons après un pronom clitique préverbal. Ces données correspondent au relevé des productions des parents de Prune au T1 (Prune, 40 mois) et de Prune au T2 (Prune, 48 mois).

Le contexte « en + X » représente le plus de collocations Mot1-Mot2 identiques entre Prune et ses parents.

	<b>Nombre de contextes de liaison entre un pronom clitique et un verbe</b>	<b>Nombre de contextes avec Mot1 identiques (en + X)</b>	<b>Nombre de collocations Mot1-Mot2 identiques (en-a)</b>
Parents de Prune (T1)	358	71	33
Prune (T2)	337	33	13

**Tableau 4 : nombre de contextes de type « en + » pour Prune et ses parents**

A la vue de ces données, le travail sur les effets d'input à travers l'étude de Mots1 identiques, voire de collocations lexicales identiques pour le développement de l'acquisition de la liaison a ses limites dans le cadre de recueils d'interactions naturelles parents-enfant. Nous constatons une grande différence d'effectifs entre le nombre de contextes en B et celui de contextes lexicaux identiques. Il conviendrait donc de réfléchir à une approche plus expérimentale pouvant contrôler davantage le lexique, agrandir le nombre de jeunes sujets et fournir des données lexicales plus importantes. Néanmoins, des travaux effectués récemment (Chabanal, 2010 ; Liégeois, Saddour et Chabanal, 2013) à partir d'un nombre réduit de contextes lexicaux présents à la fois chez les enfants et les parents attestent d'un effet de fréquence de l'input parental sur l'acquisition de patterns phonologiques. En effet, malgré ces difficultés nous avons pu, après un tri important de données et en nous focalisant sur les plus grands scores de présence des contextes chez les parents et les enfants, observer des liens entre l'input parental et le développement de la variabilité phonologique des enfants. Les rares réalisations de liaisons variables (désormais LV) par les enfants nous ont contraints à nous focaliser sur le contexte le plus productif, à savoir le contexte « Adjectif + Nom ». Dans les données ci-dessous, parmi les liaisons de type « Adjectif + Nom », nous constatons que les contextes les plus souvent réalisés justes par Prune sont également ceux qui présentent un taux de réalisation élevé dans les productions parentales.

Il existe donc un effet d'input sur les productions de Prune. La corrélation de Spearman ressort également positive et s'approche du seuil de significativité ( $\rho = 0.9$  ;  $p < 0,09$ ).

Contextes LF (Liaisons Facultatives) les plus réalisées par Prune	% de réalisations justes Prune	% de réalisations justes parents
Grand +	93% (14/15)	80% (4/5)
Petit +	91% (21/23)	77% (7/9)
Petits +	84 % (16/19)	72,7% (8/11)
Grands +	62% (5/8)	25% (1/4)
Gros +	53% (8/15)	53 % (8/15)

Tableau 5 : pourcentages de liaisons adjectif + nom réalisées justes par Prune et ses parents

## La question de l'adresse du discours et son rôle dans l'acquisition de la variation phonologique

En complément de l'analyse sur la fréquence lexicale de l'input parental, notre recherche s'intéresse aux effets de l'adresse du discours parental. On distingue différents types de discours adressés aux enfants. Il y a tout d'abord le discours adressé à tous, enfants et parents. Vosoughi, Roy, Frank et Roy (2010), le définissent comme le CAS (*Child Available Speech*), il s'agit de tout ce que peut entendre l'enfant, sans que le discours lui soit directement adressé. Il se distingue du CDS (*Child Direct Speech*) qui a lui pour caractéristique d'être un discours adressé à l'enfant mais sans que soient nécessairement précisées des informations sur l'organisation des tours de parole ou l'attention portée conjointement par l'enfant et l'adulte sur l'objet du discours (voir par exemple la méthodologie employée par Veneziano et Parisse, 2010). Ce repérage de l'organisation des tours de parole et de l'attention portée par l'adulte et l'enfant est toutefois de première importance. En effet, au cours de ces scènes d'attention conjointe, l'adulte et l'enfant portent leur attention sur l'objet, la personne ou l'événement au centre de la discussion. Ces interactions spécifiques, ponctuées de corrections ou de reformulations parentales, apportent à l'enfant des informations précieuses sur les relations entre formes et significations. Ces situations de communication entre le parent et l'enfant offrent donc un savoir beaucoup plus explicite et métalinguistique aux enfants et provoquent un effet d'input plus important que le discours adressé à l'enfant sans partage d'attention (Veneziano et Parisse, 2010).

Pour résumer, l'adresse du discours pourrait se spécifier à deux niveaux. D'une part, l'enfant aurait à sa disposition une langue orale qui ne lui serait pas adressée particulièrement, mais perceptible dans les conversations autour de lui ou adressées communément à lui et son entourage. On pourrait l'appeler « discours indirect adressé à l'enfant » ou CAS. D'autre part, l'enfant pourrait bénéficier d'un discours qui lui est spécifiquement adressé.

Pour notre étude sur la liaison et l'élision, nous avons distingué trois types de discours présents dans nos corpus denses :

- DAA : discours entre adultes (discussion entre les parents),
- DAE : discours adressé à l'enfant (similaire au CDS),
- DAT : adressé à tout le monde (le parent parle à son conjoint et son enfant en même temps. Ce type de discours est comparable au CAS).

Par ailleurs, pour prendre en compte l'importance des tours de parole et de l'attention mutuelle dans les interactions, nous avons pris soin de nous focaliser sur deux éléments dans notre relevé. Premièrement, une transcription orthographique précise nous permet de situer le contexte de la liaison, les erreurs éventuelles de l'enfant et les corrections éventuelles des parents. Deuxièmement, toutes les liaisons sont numérotées par ordre chronologique de façon à voir l'impact de la réalisation de la liaison des parents sur celles des enfants dans l'interaction.

Les résultats concernant les taux d'élision des parents selon l'adresse du discours révèlent une adaptation des parents aux performances linguistiques de leurs enfants. Ces résultats corroborent ceux précédemment obtenus concernant la variation phonologique en anglais (Foulkes, Docherty et Watt, 2005). Dans de précédents travaux (Liégeois, Saddour et Chabanal, 2012 ; 2013), nous nous sommes focalisés sur la production du schwa dans les monosyllabes : « ce, de, je, le, me, ne, que, se, te ». Les données issues de corpus denses d'interaction ont montré que le schwa était davantage réalisé en DAE, essentiellement à un stade précoce de l'acquisition (cf. Figure 1). En effet, au moment où les enfants semblent avoir acquis le phénomène en élidant dans des proportions semblables à celles de leurs parents, nous ne relevons plus de différence significative entre les taux relevés en DAE et ceux relevés en DAA. Nous pouvons en conclure que les parents de Prune et Salomé ajustent leur discours en fonction du développement linguistique de leur enfant. Nous pouvons ainsi évoquer une volonté pédagogique de la part des parents. Tous les deux semblent vouloir produire un discours surveillé tourné vers la norme orale dans le but de s'adapter aux besoins de leurs interlocuteurs suivant les différents stades de leur apprentissage.

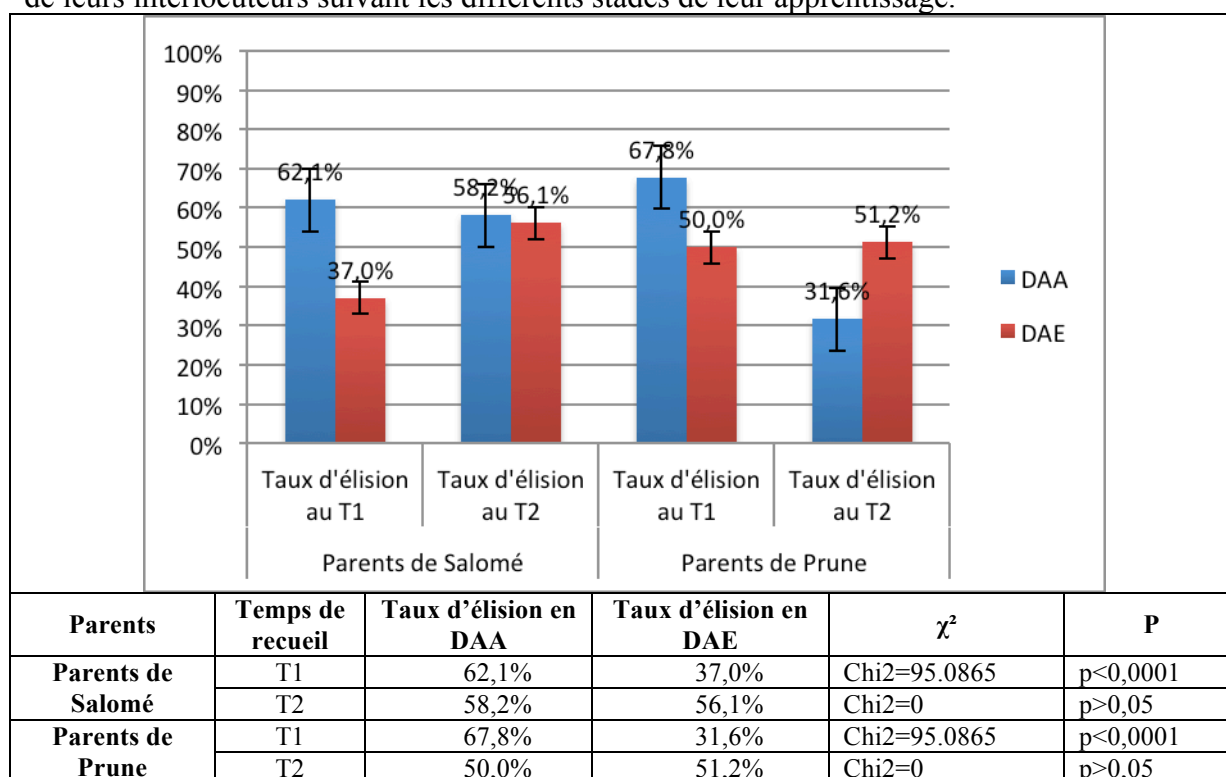


Figure 1 : taux d'élision des parents de Prune et Salomé en DAA et DAE au T1 et T2

La question de l'adaptation du parent à son interlocuteur est directement liée aux principes du courant d'interactionnisme social dont Tomasello (2003) se réclame et d'après lesquels il établit les habiletés de base pour l'acquisition du langage (attention jointe, compréhension des intentions communicatives, imitation par inversion de rôles). Selon cette théorie, les interactions sociales constituent une dynamique interindividuelle dans laquelle parents-enfant ou enseignant-apprenants exercent des influences réciproques. En clair, le bain linguistique, lieu de ces interactions, ne constitue pas simplement une influence dans un seul sens. Plutôt, une dynamique circulaire se mettrait en place : les changements linguistiques (variantes sociolinguistiques plus présentes, complexité syntaxique, variété lexicale...) dans le discours des parents/enseignants refléteraient la demande croissante de leurs enfants/apprenants à demander de nouveaux éléments pour progresser. Ces nouveaux progrès modifieraient de nouveau les attitudes des « tuteurs » qui en retour délivreraient de nouvelles données et ainsi de suite. Cette dynamique a bien été remarquée dans le cadre de nos analyses de discours parentaux.

### Le poids des contraintes linguistiques versus stylistiques sur la variation phonologique.

Cependant, plusieurs limites peuvent être soulevées quant aux résultats prenant uniquement en compte des causalités sociolinguistiques. En effet, il est clair que la réalisation du schwa est influencée par des facteurs phonologiques comme la nature du phonème qui suit directement le monosyllabe ou la longueur des mots précédant et suivant le monosyllabe (Andreassen, 2013 ; Delattre, 1951 ; Dell, 1973 ; Eychenne, 2006 ; Grammont, 1894 ; Léon, 1992 ; Racine et Grosjean, 2002). Nous avons voulu vérifier si les différences relevées dans les précédents travaux (Liégeois, Saddour et Chabanal, 2012) sont bien uniquement dues à l'adresse du discours et non à des contraintes phonologiques (type de syllabe ou type de phonème). Pour ce faire, nous avons observé, au sein du DAE et du DAA des parents de Prune, les effets possibles du contexte de droite et de gauche sur l'élision du schwa des monosyllabes. S'il s'avère que pour un même monosyllabe produit dans un contexte phonologique identique, les taux d'élision sont différents en DAE et en DAA, alors nous pourrions en conclure que seule la dimension stylistique joue un rôle. Les données que nous présentons ci-dessous sont celles extraites des énoncés produits par les parents de Prune. Les résultats font apparaître des différences nettes entre DAA et DAE imputables seulement à des causes socio-stylistiques.

Numéro du contexte	Type de mot de gauche	Structure syllabique de la syllabe de gauche	Adresse du discours	Taux d'élision (effectifs)	Structure syllabique de la syllabe de droite	Type de mot de droite
1	Début d'énoncé		DAA	90% (9/10)	CV	Monosyllabe
			DAE	34,8% (8/23)		
2	Monosyllabe	CV	DAA	89,3% (25/28)	CV	Monosyllabe
			DAE	44,2% (42/95)		
3	Monosyllabe	CV	DAA	100% (16/16)	CV	Polysyllabe
			DAE	37,5% (18/48)		

Tableau 6 : taux d'élision du schwa des monosyllabes en DAA et DAE selon les structures syllabiques environnantes

Les tests de significativité (CHI2) ayant pour but de comparer les taux de production d'élision en DAA et en DAE attestent que pour des contextes phonologiques et syllabiques identiques, les parents élident moins le schwa lorsqu'ils s'adressent à leurs enfants. Ces faits sont tendanciels pour le contexte 1 ( $\text{Chi}^2 = 6.4408$  ;  $p < 0.05$ ) et significatifs dans les deux autres (respectivement  $\text{Chi}^2 = 15.9464$  ;  $p < 0.001$  et  $\text{Chi}^2 = 16.3974$  ;  $p < 0.001$ ). La contrainte stylistique basée sur l'intention de produire un discours normé à leurs enfants semble donc bien jouer un rôle réel devant des contextes linguistiques identiques.

Ces résultats sont encourageants pour améliorer la connaissance des contraintes qui pèsent sur la variation phonologique chez les enfants. Ils sont rendus possibles par le recueil de corpus denses et les différents types de critères d'annotation choisis dans le cadre du projet ALIPE. Outre les premières conclusions qu'ils permettent, ils apportent de nouvelles voies vers l'appréhension des causalités des phénomènes variables et de leurs forces respectives dans l'acquisition des schémas d'alternances phonologiques.

## Conclusion

Tenter de comprendre l'acquisition de schémas de variation phonologique à travers la théorie basée sur l'usage a motivé un nombre important de recherches ces dernières années (Dugua, 2006 ; Nardy, 2008). La tâche nouvelle que nous nous sommes fixée autour de ce cadre a été de croiser, dans un même corpus, les productions des parents et de leurs enfants pour en mesurer les effets directs au cours du développement. Ce champ de recherche qui traite de l'effet du discours parental via la fréquence d'unités phonologiques variables au cours des interactions parents-enfant a engagé une structuration spécifique de nos données. L'approche méthodologique a permis de contribuer à étayer le scénario d'acquisition de la liaison basée sur l'usage (Chevrot, Chabanal et Dugua 2007 ; Chevrot, Dugua et Fayol, 2009 ; Chevrot, Dugua, Harnois-Delpiano, Siccardi et Spinelli, 2013). Car même si certaines de nos données (du fait du choix d'un corpus spontané) ont pu se révéler assez limitées, elles traduisent bien dans certains cas l'effet de la fréquence en input, comme nous l'avons illustré pour les contextes de liaisons variables. Cette étape est à mettre en lien avec le stade 2 du scénario d'acquisition que nous défendons. Par ailleurs, la variation phonologique étant multifactorielle (Mallet, 2008), une approche toujours plus granulaire s'est imposée à nous. Elle consiste notamment à spécifier l'adresse du discours dans l'input parental et à observer à travers la description précise des contextes phonologiques (type et nombre de syllabes, nature des phonèmes précédant et suivant la production/l'effacement du schwa...) le poids exercé par les facteurs linguistiques et sociolinguistiques.

D'une manière plus générale et pour répondre aux questions de ce numéro concernant les liens ou rapports exercés entre méthode et modèle en phonologie, il semblerait que la complémentarité soit un facteur à prendre en compte dans la réflexion. En effet, d'un point de vue méthodologique, les approches naturelles et expérimentales pourraient gagner à se compléter, car à elles seules, elles éprouvent des limites. Nous avons eu l'occasion de montrer que dans certains cas une approche écologique ne fournissait pas suffisamment de données pour valider des hypothèses sur le développement de la variation phonologique. De la même manière, la dimension expérimentale occulte la dimension interactionnelle, essentielle pour comprendre la dynamique du développement. Une des solutions envisageables serait de conduire différentes études transversales expérimentales et écologiques à des intervalles réguliers sur une période relativement longue. De cette manière nous pourrions croiser données expérimentales et écologiques afin d'obtenir d'importantes données dotées d'informations en lien avec l'input parental.

## Bibliographie

- ANDREASSEN, Hélène. (2013). *Schwa: distribution and acquisition in light of Swiss French data*. (Thèse de doctorat). University of Tromsø.
- BIERSACK, Sonja, KEMPE, Vera & KNAPTON, Lorna. (2005). Fine-Tuning Speech Registers : A Comparison of the Prosodic Features of Child-Directed and Foreigner-Directed Speech. Dans *Interspeech*, 2401–2404. Lisbonne : ISCA editors.
- BRUNER, Jerome Seymour. (1974). From communication to language: a psychological perspective. *Cognition*, 3, 255–287. <doi:10.1016/0010-0277(74)90012-2>
- BRUNER, Jerome Seymour. (1977). *The Process of Education*. Cambridge : Harvard University Press.
- BRUNER, Jerome Seymour. (1983). *Child's talk*. Cambridge : Cambridge University Press.
- CHABANAL, Damien. (2010). Effets du discours adressé à l'enfant sur l'acquisition de la liaison : étude d'un corpus dense d'une fillette de 40 mois. Actes des *Journées d'Etudes sur la Parole (JEP)*. Mons, Belgique, (cd-rom).
- CHABANAL, Damien & LIEGEOIS, Loïc (2011, juin). *Variation phonologique chez l'enfant francophone de 40 mois : développement lexical et impact du bain linguistique parental*. Communication présentée au colloque Adyloc, Paris, France.
- CHABANAL, Damien & LIEGEOIS, Loïc. (2014). Production de liaisons dans l'input parental. Dans C. Soum-Favaro, A. Coquillon & J.-P. Chevrot (dir.) *La Liaison : Approches Contemporaines*, 263-282. Berne : Peter Lang.
- CHEVROT, Jean-Pierre, CHABANAL, Damien & DUGUA, Céline. (2007). Pour un modèle de l'acquisition des liaisons basé sur l'usage: trois études de cas. *Journal of French Language Studies*, 17(1), 103–128. <doi:10.1017/S0959269506002663>
- CHEVROT, Jean-Pierre, DUGUA, Céline & FAYOL, Michel. (2009). Liaison acquisition, word segmentation and construction in French: a usage based account. *Journal of Child Language*, 36(3), 557–596. <doi:10.1017/S0305000908009124>
- CHEVROT, Jean-Pierre, DUGUA, Céline, HARNOIS-DELPINO, Mylène, SICCARDI, Anne & SPINELLI, Elsa. (2013). Liaison acquisition: debates, critical issues, future research. *Language Sciences*, 39, 83–94. doi:10.1016/j.langsci.2013.02.011
- CHOMSKY, Noam. (1965). *Aspects of the Theory of Syntax*. MIT Press. <doi:10.1016/0732-118X(86)90008-5>
- CROSS, Toni G. (1979). Mothers' speech adjustments and child language learning: Some methodological considerations. *Language Sciences*, 1(1), 3–25.
- DARWIN, Charles. (1877). A biographical sketch of an infant. *Mind*, 2(7), 285–294.
- DELATTRE, Pierre. (1951). Le jeu de l'e instable intérieur en français. *The French Review*, 24(4), 341–351.
- DELL, François. (1973). *Les règles et les sons*. Paris : Hermann.
- DEMUTH, Katherine. (2011). The Acquisition of Phonology. Dans J. Goldsmith, J. Riggle & A. C. L. Yu (dir.), *The Handbook of Phonological Theory*, 571–595. Wiley-Blackwell.
- DIESSEL, Holger. (2008). Corpus linguistics and first language acquisition. Dans A. Lüdeling & M. Kyto (dir.), *Corpus Linguistics: An International Handbook*, Vol. 2, 1197–1211. Berlin : De Gruyter.
- DILLEY, Laura C., MILLETT, Amanda L., MCAULEY, J. Delvin & BERGESON, Tonya R. (2013). Phonetic variation in consonants in infant-directed and adult-directed speech: the case of regressive place assimilation in word-final alveolar stops. *Journal of Child Language*, 41(1), 155–175. <doi:10.1017/S0305000912000670>
- DUGUA, Céline. (2006). *Liaison, segmentation lexicale et schémas syntaxiques entre 2 et 6 ans: un modèle développemental basé sur l'usage*. (Thèse de doctorat). Université Stendhal, Grenoble.

- DURAND, Jacques, LAKS, Bernard & LYCHE, Chantal. (2002). La phonologie du français contemporain: usages, variétés et structure. Dans C. Pusch & W. Raible (dir.), *Romanistische Korpuslinguistik- Korpora und gesprochene Sprache/Romance Corpus Linguistics - Corpora and Spoken Language*, 93-106. Tübingen : Gunter Narr Verlag.
- EYCHENNE, Julien. (2006). *Aspects de la phonologie du schwa dans le français contemporain*. (Thèse de doctorat). Université Toulouse-Le Mirail.
- FERGUSON, Charles A. (1964). Baby talk in six languages. *American Anthropologist*, 66, 103-114.
- FERNALD, Anne, TAESCHNER, Traute DUNN, Judy, PAPOUSEK, Mechthild, DE BOYSSON-BARDIES, Bénédicte & FUKUI, Ikuko. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, 16(3), 477-501.
- FIKKERT, Paula. (2000). Acquisition of phonology. Dans L. Cheng & R. Sysbema (dir.), *The First GLOT International State-of-the Article Book. The Latest in Linguistics*, 221-250. Berlin, Boston : De Gruyter Mouton.
- FOULKES, Paul, DOCHERTY, Gerard & WATT, Dominic. (2005). Phonological Variation in Child-Directed Speech. *Language*, 81(1), 177-206. <doi:10.1353/lan.2005.0018>
- GARNICA, Olga. K. (1977). Some prosodic and paralinguistic features of speech to young children. Dans C. E. Snow & C. A. Ferguson (dir.), *Talking to Children Language Input and Acquisition*, 63-88. Cambridge : Cambridge University Press.
- GRAMMONT, Maurice. (1894). *Le patois de la Franche-Montagne et en particulier de Damprichard (Franche-Comté). IV : La loi des trois consonnes*.
- GREGOIRE, Antoine. (1948). L'apprentissage du langage. *Lingua*, 1, 162-174.
- HALLIDAY, Michael. (1973). *Explorations in the functions of language*. Londres : Edward Arnold.
- INGRAM, David. (1989). *First Language Acquisition: Method, Description and Explanation*. Cambridge : Cambridge University Press.
- LEON, Pierre. (1992). *Phonétisme et Prononciations du français*. Paris : Nathan-Fac.
- LEOPOLD, Werner. (1939). Speech development of a bilingual child. *A linguist's record*. Vol.1. Evanston : Northwestern University Press.
- LIÉGEOIS, Loïc, CHABANAL, Damien & CHANIER, Thierry. (2011). La liaison en discours adressé à l'enfant, spécificités et impacts sur l'acquisition. Communication présentée au colloque du Réseau Français de Phonologie, 1-3 Juillet 2011, Tours.
- LIEGEOIS, Loïc, CHANIER, Thierry & CHABANAL, Damien. (2014). *Corpus globaux ALIPE : Interactions parents-enfant annotées pour l'étude de la liaison*. Nancy : Ortolang. <<http://hdl.handle.net/11041/alipe-000853>>
- LIEGEOIS, Loïc, SADDOUR, Inès & CHABANAL, Damien. (2012). L'élision du schwa dans les interactions parents-enfant : étude de corpus. Actes des Journées d'Études sur la Parole, colloque international, 4-8 juin 2012, Grenoble, 313-320. <<http://hal.archives-ouvertes.fr/hal-00707773>>
- LIEGEOIS, Loïc, SADDOUR, Inès & CHABANAL, Damien. (2013). Input characteristics and impact on the acquisition of phonological variables in French. Communication présentée au Child Language Seminar, 24-25 juin 2013, Manchester.
- LIEGEOIS, Loïc. (2013). De l'analyse au partage des données, quel(s) format(s) choisir ? L'exemple d'un corpus d'interactions parents-enfant. Dans M. Damiani, K. Dolar, C. Florez-Pulido, R. Loth, J. Magnier & A. Pegaz (dir.) *Traitement de corpus (Actes de Coldoc 2012)*, 128-142. Paris : Modyco. <<http://hal.archives-ouvertes.fr/hal-00850172>>
- LIEGEOIS, Loïc. (2014). *Usage des variables phonologiques dans un corpus d'interactions naturelles parents-enfant : impact du bain linguistique et dispositifs cognitifs d'apprentissage*. (Thèse de doctorat). Université Blaise Pascal, Clermont Université.



- LIEVEN, Elena & BERHENS, Heike. (2012). Dense Sampling. Dans E. Hoff (Ed.), *Research Methods in Child Language. A Practical Guide*, 226–239. Wiley-Blackwell.
- LOISEAU, Sylvain. (2011). Les faits statistiques comme objectivation ou comme interprétation : statistiques et modèles basés sur l'usage. *Travaux de linguistique*, 62, 59–78.
- MACWHINNEY, Brian & SNOW, Catherine E. (1985). The child language data exchange system. *Journal of Child Language*, 12(2), 271–295.
- MACWHINNEY, Brian. (2000). *The CHILDES Project: Tools for Analyzing Talk*. Mahwah : Lawrence Erlbaum Associates.
- MALLET, Géraldine. (2008). *La liaison en français : descriptions et analyses dans le corpus PFC*. (Thèse de Doctorat). Université Paris Ouest Nanterre la Défense.
- MORGENSTERN, Aliyah & PARISSE, Christophe. (2007). Codage et interprétation du langage spontané d'enfants de 1 à 3 ans. *Corpus*, (6), 55–78.
- NARDY, Aurélie. (2008). *Acquisition des variables sociolinguistiques entre 2 et 6 ans : facteurs sociologiques et influences des interactions au sein du réseau social* (Thèse de doctorat). Université Stendhal, Grenoble.
- RACINE, Isabelle & GROSJEAN, François. (2002). La production du E caduc facultatif est-elle prévisible ? Un début de réponse. *Journal of French Language Studies*, 12(03), 307–326. <doi:10.1017/S0959269502000340>
- ROWLAND, Caroline, FLETCHER, Sarah L. & FREUDENTHAL, Daniel. (2008). How big is big enough? Assessing the reliability of data from naturalistic samples. Dans H. Behrens (Ed.), *Corpora in Language Acquisition Research. History, methods, perspectives*, 1–24. John Benjamins Publishing Company.
- SCHAEFFLER, Sonja & KEMPE, Vera. (2007). Mothers Are Less Efficient in Employing Prosodic Disambiguation in Child-Directed Speech than Non-Mothers: Is There a Trade-Off Between Affective and Linguistic Prosody? Dans J. Trouvain & J. B. Barry (dir.), *Proceedings of the 16th International Congress of Phonetic Sciences*, 2109–2112.
- SNOW, Catherine. E. (1972). Mothers' speech to children learning language. *Child Development*, 43(2), 549–565.
- TAINÉ, Hippolyte. (1876). L'acquisition du langage chez les enfants et dans l'espèce humaine. *Revue Philosophique de La France et de l'Etranger*, 1, 357–395.
- TAINÉ, Hippolyte. (1877). M. Taine on the Acquisition of Language by Children. *Mind*, 2(6), 252–259.
- TOMASELLO, Michael & STAHL, Daniel. (2004). Sampling childrens spontaneous speech : how much is enough ? *Journal of Child Language*, 31(1), 101–121.
- TOMASELLO, Michael. (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge : Harvard University Press.
- VENEZIANO, Eddy & PARISSE, Christophe. (2010). The acquisition of early verbs in French: Assessing the role of conversation and of child-directed speech. *First Language*, 30(3-4), 287–311.
- VOSOUGHI, Soroush, ROY, Brandon. C., FRANCK, Michael. C. & ROY, Deb. (2010). Contributions of Prosodic and Distributional Features of Caregivers' Speech in Early Word Learning. *32nd Annual Cognitive Science Conference*. Portland.
- WATT, Dominic, DOCHERTY, Gerard J. & FOULKES, Paul. (2003). First accent acquisition: a study of phonetic variation in child-directed speech. Dans *Proceedings of the 15th International Congress of Phonetic Sciences*, 1959–1962. Barcelona.
- WAUQUIER-GRAVELINES, Sophie. (2005). Acquisition et développement phonologiques. Dans N. Nguyen, S. Wauquier-Gravelines & J. Durand (dir.), *Phonologie et phonétique, forme et substance*, 325–345. Paris : Hermès.